Miskolc Mathematical Notes
Vol. 17 (2016), No. 1, pp. 151–169
HU e-ISSN 1787-2413
DOI: 10.18514/MMN.2016.1794

INTEGRAL INEQUALITIES FOR CONVEX FUNCTIONS AND APPLICATIONS FOR DIVERGENCE MEASURES

S. S. DRAGOMIR

Received 07 October, 2015

Abstract. Some new integral inequalities for convex functions with applications for f-divergence measures in information theory are given.

2010 *Mathematics Subject Classification:* 26D15; 26D20; 94A05

Keywords: convex functions, integral inequalities, Jensen's type inequalities, Fejér type inequalities, Lebesgue integral, divergence measures, f-divergence

1. INTRODUCTION

Suppose that I is an interval of real numbers with interior $\mathring{I}$ and $\Phi : I \to \mathbb{R}$ is a convex function on I. Then Φ is continuous on $\mathring{I}$ and has finite left and right derivatives at each point of $\mathring{I}$. Moreover, if $x, y \in \mathring{I}$ and $x < y$, then $\Phi'_-(x) \le \Phi'_+(x) \le \Phi'_-(y) \le \Phi'_+(y)$ which shows that both Φ'_- and Φ'_+ are nondecreasing function on $\mathring{I}$. It is also known that a convex function must be differentiable except for at most countably many points.

For a convex function $\Phi : I \to \mathbb{R}$, the subdifferential of Φ denoted by $\partial\Phi$ is the set of all functions $\varphi : I \to [-\infty, \infty]$ such that $\varphi\left(\mathring{I}\right) \subset \mathbb{R}$ and

$$\Phi(x) \ge \Phi(a) + (x-a)\varphi(a) \text{ for any } x, a \in I. \tag{1.1}$$

It is also well known that if Φ is convex on I, then $\partial\Phi$ is nonempty, $\Phi'_-, \Phi'_+ \in \partial\Phi$ and if $\varphi \in \partial\Phi$, then

$$\Phi'_-(x) \le \varphi(x) \le \Phi'_+(x) \text{ for any } x \in \mathring{I}.$$

In particular, φ is a nondecreasing function.

If Φ is differentiable and convex on $\mathring{I}$, then $\partial\Phi = \{\Phi'\}$.

Let $(\Omega, \mathcal{A}, \mu)$ be a measurable space consisting of a set Ω, a $\sigma-$algebra $\mathcal{A}$ of parts of Ω and a countably additive and positive measure μ on $\mathcal{A}$ with values in $\mathbb{R} \cup \{\infty\}$. For a $\mu-$measurable function $w : \Omega \to \mathbb{R}$, with $w(x) \ge 0$ for μ – a.e.(almost every)

$x \in \Omega$, consider the Lebesgue space

$$L_w(\Omega,\mu) := \{f : \Omega \to \mathbb{R},\ f \text{ is } \mu\text{-measurable and } \int_\Omega |f(x)| w(x) d\mu(x) < \infty\}.$$

For simplicity of notation we write everywhere in the sequel $\int_\Omega w d\mu$ instead of $\int_\Omega w(x) d\mu(x)$.

In order to provide a reverse of the celebrated Jensen's integral inequality for convex functions, S. S. Dragomir obtained in 2002 [15] the following result:

Theorem 1. *Let $\Phi : [m,M] \subset \mathbb{R} \to \mathbb{R}$ be a differentiable convex function on (m,M) and $f : \Omega \to [m,M]$ so that $\Phi \circ f,\ f,\ \Phi' \circ f,\ (\Phi' \circ f) f \in L_w(\Omega,\mu)$, where $w \geq 0$ μ-a.e. (almost everywhere) on Ω with $\int_\Omega w d\mu = 1$. Then we have the inequality:*

$$\begin{aligned}
0 &\leq \int_\Omega (\Phi \circ f) w d\mu - \Phi\left(\int_\Omega f w d\mu\right) \\
&\leq \int_\Omega (\Phi' \circ f) f w d\mu - \int_\Omega (\Phi' \circ f) w d\mu \int_\Omega f w d\mu \\
&\leq \frac{1}{2}\left[\Phi'_-(M) - \Phi'_-(m)\right] \int_\Omega \left| f - \int_\Omega f w d\mu \right| w d\mu.
\end{aligned} \tag{1.2}$$

Corollary 1. *Let $\Phi : [m,M] \to \mathbb{R}$ be a differentiable convex function on (m,M). If $x_i \in [m,M]$ and $w_i \geq 0$ $(i = 1,\dots,n)$ with $W_n := \sum_{i=1}^n w_i = 1$, then one has the counterpart of Jensen's weighted discrete inequality:*

$$\begin{aligned}
0 &\leq \sum_{i=1}^n w_i \Phi(x_i) - \Phi\left(\sum_{i=1}^n w_i x_i\right) \\
&\leq \sum_{i=1}^n w_i \Phi'(x_i) x_i - \sum_{i=1}^n w_i \Phi'(x_i) \sum_{i=1}^n w_i x_i \\
&\leq \frac{1}{2}\left[\Phi'_+(M) - \Phi'_-(m)\right] \sum_{i=1}^n w_i \left| x_i - \sum_{j=1}^n w_j x_j \right|.
\end{aligned} \tag{1.3}$$

Remark 1. We notice that the inequality between the first and the second term in (1.3) was proved in 1994 by Dragomir & Ionescu, see [17].

If $f,g : \Omega \to \mathbb{R}$ are $\mu-$measurable functions and $f,g,fg \in L_w(\Omega,\mu)$, then we may consider the *Čebyšev functional*

$$T_w(f,g) := \int_\Omega f g w d\mu - \int_\Omega f w d\mu \int_\Omega g w d\mu. \tag{1.4}$$

The following result is known in the literature as the *Grüss inequality*

$$|T_w(f,g)| \leq \frac{1}{4}(\Gamma - \gamma)(\Delta - \delta), \tag{1.5}$$

provided

$$-\infty < \gamma \leq f(x) \leq \Gamma < \infty, \quad -\infty < \delta \leq g(x) \leq \Delta < \infty \tag{1.6}$$

for μ – a.e. $x \in \Omega$.

The constant $\frac{1}{4}$ is sharp in the sense that it cannot be replaced by a smaller quantity.

If we assume that $-\infty < \gamma \leq f(x) \leq \Gamma < \infty$ for μ – a.e. $x \in \Omega$, then by the Grüss inequality for $g = f$ and by the Schwarz's integral inequality, we have

$$\int_\Omega \left| f - \int_\Omega f w d\mu \right| w d\mu \leq \left[\int_\Omega f^2 w d\mu - \left(\int_\Omega f w d\mu \right)^2 \right]^{\frac{1}{2}} \leq \frac{1}{2} (\Gamma - \gamma). \tag{1.7}$$

On making use of the results (1.2) and (1.7), we can state the following string of reverse inequalities

$$\begin{aligned} 0 &\leq \int_\Omega (\Phi \circ f) w d\mu - \Phi\left(\int_\Omega f w d\mu \right) \\ &\leq \int_\Omega (\Phi' \circ f) f w d\mu - \int_\Omega (\Phi' \circ f) w d\mu \int_\Omega f w d\mu \\ &\leq \frac{1}{2} \left[\Phi'_+(M) - \Phi'_-(m) \right] \int_\Omega \left| f - \int_\Omega w f d\mu \right| w d\mu \\ &\leq \frac{1}{2} \left[\Phi'_+(M) - \Phi'_-(m) \right] \left[\int_\Omega f^2 w d\mu - \left(\int_\Omega f w d\mu \right)^2 \right]^{\frac{1}{2}} \\ &\leq \frac{1}{4} \left[\Phi'_+(M) - \Phi'_-(m) \right] (M - m), \end{aligned} \tag{1.8}$$

provided that $\Phi : [m, M] \subset \mathbb{R} \to \mathbb{R}$ is a differentiable convex function on (m, M) and $f : \Omega \to [m, M]$ so that $\Phi \circ f$, f, $\Phi' \circ f$, $(\Phi' \circ f) f \in L_w(\Omega, \mu)$, where $w \geq 0$ μ-a.e. on Ω with $\int_\Omega w d\mu = 1$.

Remark 2. We notice that the inequality between the first, second and last term from (1.8) was proved in the general case of positive linear functionals in 2001 by S.S. Dragomir in [14].

For recent inequalities for convex functions, see [1, 2, 4, 10, 18, 23, 29, 33–36, 38, 43–45] and [51].

Motivated by the above results, we establish in this paper some integral inequalities in which we provide upper and lower bounds for the quantity $\int_\Omega (\Phi \circ f) w d\mu$ and obtain some generalization for the celebrated Fejér's inequality [19]. Applications for divergence measures in information theory are provided as well.

2. Main results

The following result holds:

Theorem 2. *Let $\Phi : [m,M] \subset \mathbb{R} \to \mathbb{R}$ be a convex function on $[m,M]$ and $f : \Omega \to [m,M]$ so that $\Phi \circ f,\ f \in L_w(\Omega,\mu)$, where $w \geq 0$ μ-a.e. on Ω with $\int_\Omega w\,d\mu = 1$. Then we have the inequalities:*

$$
\begin{aligned}
&\Phi\left(\frac{m+M}{2}\right) + \varphi\left(\frac{m+M}{2}\right)\int_\Omega \left(f - \frac{m+M}{2}\right) w\,d\mu \\
&\leq \int_\Omega (\Phi \circ f)\, w\,d\mu \\
&\leq \frac{\Phi(m)+\Phi(M)}{2} + \frac{\Phi(M)-\Phi(m)}{M-m}\int_\Omega \left(f - \frac{m+M}{2}\right) w\,d\mu,
\end{aligned}
\tag{2.1}
$$

where $\varphi\left(\frac{m+M}{2}\right) \in \left[\Phi'_-\left(\frac{m+M}{2}\right), \Phi'_+\left(\frac{m+M}{2}\right)\right]$.

Proof. By the gradient inequality (1.1) we have

$$
\Phi(t) - \Phi\left(\frac{m+M}{2}\right) \geq \left(t - \frac{m+M}{2}\right)\varphi\left(\frac{m+M}{2}\right) \tag{2.2}
$$

where $\varphi\left(\frac{m+M}{2}\right) \in \left[\Phi'_-\left(\frac{m+M}{2}\right), \Phi'_+\left(\frac{m+M}{2}\right)\right]$ for any $t \in [m,M]$.

This inequality implies that

$$
\Phi(f(x)) \geq \Phi\left(\frac{m+M}{2}\right) + \left(f(x) - \frac{m+M}{2}\right)\varphi\left(\frac{m+M}{2}\right) \tag{2.3}
$$

for any $x \in \Omega$.

If we multiply (2.3) by $w \geq 0$ μ-a.e and integrate on Ω, we get the first inequality in (2.1).

By the convexity of Φ we also have

$$
\begin{aligned}
\Phi(t) &= \Phi\left(\frac{M-t}{M-m}m + \frac{t-m}{M-m}M\right) \\
&\leq \frac{M-t}{M-m}\Phi(m) + \frac{t-m}{M-m}\Phi(M) \\
&= \frac{\Phi(m)+\Phi(M)}{2} + \left(\frac{M-t}{M-m} - \frac{1}{2}\right)\Phi(m) + \left(\frac{t-m}{M-m} - \frac{1}{2}\right)\Phi(M) \\
&= \frac{\Phi(m)+\Phi(M)}{2} - \Phi(m)\left(\frac{t - \frac{m+M}{2}}{M-m}\right) + \Phi(M)\left(\frac{t - \frac{m+M}{2}}{M-m}\right) \\
&= \frac{\Phi(m)+\Phi(M)}{2} + \frac{\Phi(M)-\Phi(m)}{M-m}\left(t - \frac{m+M}{2}\right)
\end{aligned}
$$

for any $t \in [m,M]$.

This inequality implies that

$$
\Phi(f(x)) \leq \frac{\Phi(m)+\Phi(M)}{2} + \frac{\Phi(M)-\Phi(m)}{M-m}\left(f(x) - \frac{m+M}{2}\right) \tag{2.4}
$$

for any $x \in \Omega$.

If we multiply (2.4) by $w \geq 0$ μ-a.e and integrate on Ω, we get the second inequality in (2.1). $\square$

Corollary 2. *With the assumptions of Theorem 2 and if*

$$\int_\Omega \left(f - \frac{m+M}{2} \right) w d\mu = 0, \tag{2.5}$$

then we have

$$\Phi\left(\frac{m+M}{2}\right) \leq \int_\Omega (\Phi \circ f)\, w d\mu \leq \frac{\Phi(m) + \Phi(M)}{2}. \tag{2.6}$$

Remark 3. With the assumptions of Theorem 2 and if

$$0 \in \left[\Phi'_-\left(\frac{m+M}{2}\right), \Phi'_+\left(\frac{m+M}{2}\right) \right],$$

then the first inequality in (2.6) is valid.

If either

$$\Phi(M) \leq \Phi(m) \text{ and } \int_\Omega \left(f - \frac{m+M}{2} \right) w d\mu \geq 0$$

or

$$\Phi(M) \geq \Phi(m) \text{ and } \int_\Omega \left(f - \frac{m+M}{2} \right) w d\mu \leq 0,$$

then the second inequality in (2.6) is valid.

The following result also holds:

Theorem 3. *Let $\Phi : [m, M] \subset \mathbb{R} \to \mathbb{R}$ be a convex function on $[m, M]$ and $f : \Omega \to [m, M]$ so that $\Phi \circ f$, f, $\left(f - \frac{m+M}{2}\right)(\Phi' \circ f) \in L_w(\Omega, \mu)$, where $w \geq 0$ μ-a.e. on Ω with $\int_\Omega w d\mu = 1$. Then we have the inequalities:*

$$\begin{aligned}
&\frac{2}{M-m} \int_m^M \Phi(s)\, ds - \frac{\Phi(M) + \Phi(m)}{2} + \frac{\Phi(M) - \Phi(m)}{M-m} \int_\Omega \left(f - \frac{m+M}{2} \right) w d\mu \\
&\leq \int_\Omega (\Phi \circ f)\, w d\mu \\
&\leq \frac{1}{M-m} \int_m^M \Phi(t)\, dt + \int_\Omega (\Phi' \circ f)\, w d\mu \int_\Omega \left(f - \frac{m+M}{2} \right) w d\mu \\
&\quad + \int_\Omega \left[(\Phi' \circ f) - \int_\Omega (\Phi' \circ f)\, w d\mu \right] f w d\mu.
\end{aligned} \tag{2.7}$$

Proof. By the gradient inequality (1.1) we have

$$\Phi(t) - \Phi(s) \geq (t - s)\, \Phi'(s), \tag{2.8}$$

for any $t \in [m, M]$ and almost any $s \in [m, M]$.

Integrating over s on $[m,M]$ and dividing by $M-m$ we get

$$\begin{aligned}
\Phi(t) &\geq \frac{1}{M-m}\int_m^M \Phi(s)\,ds + \frac{1}{M-m}\int_m^M (t-s)\,\Phi'(s)\,ds \qquad (2.9)\\
&= \frac{1}{M-m}\int_m^M \Phi(s)\,ds + \frac{1}{M-m}\left[(t-s)\,\Phi(s)|_m^M + \int_m^M \Phi(s)\,ds\right]\\
&= \frac{1}{M-m}\int_m^M \Phi(s)\,ds\\
&\quad + \frac{1}{M-m}\left[\int_m^M \Phi(s)\,ds - (M-t)\,\Phi(M) - (t-m)\,\Phi(m)\right]\\
&= \frac{2}{M-m}\int_m^M \Phi(s)\,ds - \frac{(M-t)\,\Phi(M)+(t-m)\,\Phi(m)}{M-m}
\end{aligned}$$

for any $t\in[m,M]$.

Observe that

$$\begin{aligned}
&\frac{(M-t)\,\Phi(M)+(t-m)\,\Phi(m)}{M-m} - \frac{\Phi(M)+\Phi(m)}{2} \qquad (2.10)\\
&= \left(\frac{M-t}{M-m}-\frac{1}{2}\right)\Phi(M) + \left(\frac{t-m}{M-m}-\frac{1}{2}\right)\Phi(m)\\
&= \frac{\Phi(M)-\Phi(m)}{M-m}\left(\frac{m+M}{2}-t\right)
\end{aligned}$$

for any $t\in[m,M]$.

Then we have

$$\begin{aligned}
&\frac{2}{M-m}\int_m^M \Phi(s)\,ds - \frac{(M-t)\,\Phi(M)+(t-m)\,\Phi(m)}{M-m} \qquad (2.11)\\
&= \frac{2}{M-m}\int_m^M \Phi(s)\,ds - \frac{(M-t)\,\Phi(M)+(t-m)\,\Phi(m)}{M-m}\\
&\quad + \frac{\Phi(M)+\Phi(m)}{2} - \frac{\Phi(M)+\Phi(m)}{2}\\
&= \frac{2}{M-m}\int_m^M \Phi(s)\,ds + \frac{\Phi(M)-\Phi(m)}{M-m}\left(t-\frac{m+M}{2}\right)\\
&\quad - \frac{\Phi(M)+\Phi(m)}{2}
\end{aligned}$$

and by (2.9) we have

$$\Phi(t) \geq \frac{2}{M-m}\int_m^M \Phi(s)\,ds + \frac{\Phi(M)-\Phi(m)}{M-m}\left(t-\frac{m+M}{2}\right) - \frac{\Phi(M)+\Phi(m)}{2}$$

for any $t\in[m,M]$.

This inequality implies that

$$\Phi(f(x)) \geq \frac{2}{M-m}\int_m^M \Phi(s)\,ds + \frac{\Phi(M)-\Phi(m)}{M-m}\left(f(x)-\frac{m+M}{2}\right) - \frac{\Phi(M)+\Phi(m)}{2} \tag{2.12}$$

for any $x \in \Omega$.

If we multiply (2.12) by $w \geq 0$ μ-a.e and integrate on Ω, we get the first inequality in (2.7).

Integrating over t in (2.8) and dividing by $M-m$ we get

$$\frac{1}{M-m}\int_m^M \Phi(t)\,dt - \Phi(s) \geq \left(\frac{m+M}{2}-s\right)\Phi'(s),$$

for almost every $s \in [m,M]$, which is equivalent to

$$\frac{1}{M-m}\int_m^M \Phi(t)\,dt + \left(s-\frac{m+M}{2}\right)\Phi'(s) \geq \Phi(s)$$

for almost every $s \in [m,M]$.

This inequality implies that

$$\frac{1}{M-m}\int_m^M \Phi(t)\,dt + \left(f(x)-\frac{m+M}{2}\right)\Phi'(f(x)) \geq \Phi(f(x)) \tag{2.13}$$

for μ–a.e. $x \in \Omega$.

If we multiply (2.13) by $w \geq 0$ μ-a.e and integrate on Ω, we get the inequality

$$\int_\Omega (\Phi\circ f)\,w\,d\mu \leq \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \int_\Omega \left(f-\frac{m+M}{2}\right)(\Phi'\circ f)\,w\,d\mu.$$

Now, since a simple calculation reveals that

$$\begin{aligned}
&\frac{1}{M-m}\int_m^M \Phi(t)\,dt + \int_\Omega \left(f-\frac{m+M}{2}\right)(\Phi'\circ f)\,w\,d\mu \\
&= \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \int_\Omega (\Phi'\circ f)\,w\,d\mu \int_\Omega \left(f-\frac{m+M}{2}\right)w\,d\mu \\
&\quad + \int_\Omega \left(f-\frac{m+M}{2}\right)\left[(\Phi'\circ f) - \int_\Omega (\Phi'\circ f)\,w\,d\mu\right]w\,d\mu \\
&= \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \int_\Omega \left[(\Phi'\circ f) - \int_\Omega (\Phi'\circ f)\,w\,d\mu\right] f\,w\,d\mu,
\end{aligned}$$

and the second part of (2.7) is also proved. □

Remark 4. Making use of (1.8), we have the following string of inequalities

$$\int_\Omega (\Phi\circ f)\,w\,d\mu \tag{2.14}$$

$$
\begin{aligned}
&\le \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \int_\Omega (\Phi'\circ f)\,w\,d\mu \int_\Omega \left(f - \frac{m+M}{2}\right) w\,d\mu \\
&\quad + \int_\Omega \left[(\Phi'\circ f) - \int_\Omega (\Phi'\circ f)\,w\,d\mu\right] f\,w\,d\mu \\
&\le \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \int_\Omega (\Phi'\circ f)\,w\,d\mu \int_\Omega \left(f - \frac{m+M}{2}\right) w\,d\mu \\
&\quad + \frac{1}{2}\left[\Phi'_+(M) - \Phi'_-(m)\right] \int_\Omega \left| f - \int_\Omega f\,w\,d\mu \right| w\,d\mu \\
&\le \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \int_\Omega (\Phi'\circ f)\,w\,d\mu \int_\Omega \left(f - \frac{m+M}{2}\right) w\,d\mu \\
&\quad + \frac{1}{2}\left[\Phi'_+(M) - \Phi'_-(m)\right] \left[\int_\Omega f^2 w\,d\mu - \left(\int_\Omega f\,w\,d\mu\right)^2\right]^{\frac{1}{2}} \\
&\le \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \int_\Omega (\Phi'\circ f)\,w\,d\mu \int_\Omega \left(f - \frac{m+M}{2}\right) w\,d\mu \\
&\quad + \frac{1}{4}\left[\Phi'_+(M) - \Phi'_-(m)\right](M-m).
\end{aligned}
$$

Corollary 3. *With the assumptions of Theorem 3 and if condition (2.5) holds, then*

$$
\begin{aligned}
&\frac{2}{M-m}\int_m^M \Phi(s)\,ds - \frac{\Phi(M)+\Phi(m)}{2} \qquad (2.15)\\
&\le \int_\Omega (\Phi\circ f)\,w\,d\mu \\
&\le \frac{1}{M-m}\int_m^M \Phi(s)\,ds + \int_\Omega \left[(\Phi'\circ f) - \int_\Omega (\Phi'\circ f)\,w\,d\mu\right] f\,w\,d\mu \\
&\le \frac{1}{M-m}\int_m^M \Phi(s)\,ds + \frac{1}{2}\left[\Phi'_+(M) - \Phi'_-(m)\right] \int_\Omega \left| f - \int_\Omega f\,w\,d\mu \right| w\,d\mu \\
&\le \frac{1}{M-m}\int_m^M \Phi(s)\,ds + \frac{1}{2}\left[\Phi'_+(M) - \Phi'_-(m)\right] \times \left[\int_\Omega f^2 w\,d\mu - \left(\int_\Omega f\,w\,d\mu\right)^2\right]^{\frac{1}{2}} \\
&\le \frac{1}{M-m}\int_m^M \Phi(s)\,ds + \frac{1}{4}\left[\Phi'_+(M) - \Phi'_-(m)\right](M-m).
\end{aligned}
$$

The case of functions of real variable is of interest due to its connection to *Fejér's inequality* that states that [19]

$$
\Phi\left(\frac{a+b}{2}\right)\int_a^b w(t)\,dt \le \int_a^b \Phi(t)\,w(t)\,dt \le \frac{\Phi(a)+\Phi(b)}{2}\int_a^b w(t)\,dt, \qquad (2.16)
$$

provided $\Phi : [a,b] \to \mathbb{R}$ is convex and $w : [a,b] \to [0,\infty)$ is integrable and symmetric on $[a,b]$, i.e. $w(a+b-t) = w(t)$ for all $t \in [a,b]$.

If $\Phi : [a,b] \to \mathbb{R}$ is convex and $w : [a,b] \to [0,\infty)$ is integrable with $\int_a^b w(t)\,dt = 1$, then from (2.1) we have

$$\begin{aligned} &\Phi\left(\frac{a+b}{2}\right) + \varphi\left(\frac{a+b}{2}\right)\int_a^b \left(t - \frac{a+b}{2}\right) w(t)\,dt \\ &\le \int_a^b \Phi(t)\,w(t)\,dt \\ &\le \frac{\Phi(a)+\Phi(b)}{2} + \frac{\Phi(b)-\Phi(a)}{b-a}\int_a^b \left(t - \frac{a+b}{2}\right) w(t)\,dt, \end{aligned} \tag{2.17}$$

where $\varphi\left(\frac{a+b}{2}\right) \in \left[\Phi'_-\left(\frac{a+b}{2}\right), \Phi'_+\left(\frac{a+b}{2}\right)\right]$ and

$$\begin{aligned} &\frac{2}{b-a}\int_a^b \Phi(s)\,ds - \frac{\Phi(b)+\Phi(a)}{2} \\ &\quad + \frac{\Phi(b)-\Phi(a)}{b-a}\int_a^b \left(t - \frac{a+b}{2}\right) w(t)\,dt \\ &\le \int_a^b \Phi(t)\,w(t)\,dt \\ &\le \frac{1}{b-a}\int_a^b \Phi(t)\,dt + \int_a^b \Phi'(t)\,w(t)\,dt \int_a^b \left(t - \frac{a+b}{2}\right) w(t)\,dt \\ &\quad + \int_a^b \left[\Phi'(t) - \int_a^b \left(\Phi'(s)\right) w(s)\,ds\right] t\,w(t)\,dt \\ &\le \frac{1}{b-a}\int_a^b \Phi(t)\,dt + \int_a^b \Phi'(t)\,w(t)\,dt \int_a^b \left(t - \frac{a+b}{2}\right) w(t)\,dt \\ &\quad + \frac{1}{2}\left[\Phi'_+(b) - \Phi'_-(a)\right]\int_a^b \left|t - \int_a^b s\,w(s)\,ds\right| w(t)\,dt \\ &\le \frac{1}{b-a}\int_a^b \Phi(t)\,dt + \int_a^b \Phi'(t)\,w(t)\,dt \int_a^b \left(t - \frac{a+b}{2}\right) w(t)\,dt \\ &\quad + \frac{1}{2}\left[\Phi'_+(b) - \Phi'_-(a)\right]\left[\int_a^b t^2 w(t)\,dt - \left(\int_a^b t\,w(t)\,dt\right)^2\right]^{\frac{1}{2}} \\ &\le \frac{1}{b-a}\int_a^b \Phi(t)\,dt + \int_a^b \Phi'(t)\,w(t)\,dt \int_a^b \left(t - \frac{a+b}{2}\right) w(t)\,dt \end{aligned} \tag{2.18}$$

$$+\frac{1}{2}\left[\Phi'_{+}(b)-\Phi'_{-}(a)\right](b-a).$$

If $w:[a,b]\to[0,\infty)$ is integrable and

$$\int_a^b \left(t-\frac{a+b}{2}\right)w(t)\,dt=0, \tag{2.19}$$

then from (2.17) we have the Fejér's inequality (2.16) for $\int_a^b w(t)\,dt=1$. We observe that the condition (2.19) is more general than the symmetry of the functions w. If w is symmetric, then the function $h(t):=\left(t-\frac{a+b}{2}\right)w(t)$ is antisymmetric and then $\int_a^b h(t)\,dt=0$.

If (2.19) is satisfied, then from (2.18) we get

$$\begin{aligned}
&\frac{2}{b-a}\int_a^b \Phi(s)\,ds-\frac{\Phi(b)+\Phi(a)}{2} \qquad (2.20)\\
&\le \int_a^b \Phi(t)\,w(t)\,dt\\
&\le \frac{1}{b-a}\int_a^b \Phi(t)\,dt+\int_a^b\left[\Phi'(t)-\int_a^b (\Phi'(s))\,w(s)\,ds\right]t\,w(t)\,dt\\
&\le \frac{1}{b-a}\int_a^b \Phi(t)\,dt+\frac{1}{2}\left[\Phi'_{+}(b)-\Phi'_{-}(a)\right]\int_a^b\left|t-\int_a^b s\,w(s)\,ds\right|w(t)\,dt\\
&\le \frac{1}{b-a}\int_a^b \Phi(t)\,dt+\frac{1}{2}\left[\Phi'_{+}(b)-\Phi'_{-}(a)\right]\times\left[\int_a^b t^2w(t)\,dt-\left(\int_a^b t\,w(t)\,dt\right)^2\right]^{\frac{1}{2}}\\
&\le \frac{1}{b-a}\int_a^b \Phi(t)\,dt+\frac{1}{4}\left[\Phi'_{+}(b)-\Phi'_{-}(a)\right](b-a).
\end{aligned}$$

Now, if we take in (2.17) $w(t)=\frac{1}{b-a}, t\in[a,b]$ then we get the Hermite-Hadamard inequality

$$\Phi\left(\frac{a+b}{2}\right)\le\frac{1}{b-a}\int_a^b \Phi(t)\,dt\le\frac{\Phi(a)+\Phi(b)}{2}. \tag{2.21}$$

The case of discrete measure produces the following inequalities of interest for a convex function $\Phi:[m,M]\to\mathbb{R}$,

$$\begin{aligned}
&\Phi\left(\frac{m+M}{2}\right)+\varphi\left(\frac{m+M}{2}\right)\sum_{i=1}^n w_i\left(x_i-\frac{m+M}{2}\right) \qquad (2.22)\\
&\le \sum_{i=1}^n w_i\,\Phi(x_i)
\end{aligned}$$

$$\leq \frac{\Phi(m)+\Phi(M)}{2} + \frac{\Phi(M)-\Phi(m)}{M-m} \sum_{i=1}^{n} w_i \left(x_i - \frac{m+M}{2} \right),$$

where $\varphi\left(\frac{m+M}{2}\right) \in \left[\Phi'_{-}\left(\frac{m+M}{2}\right), \Phi'_{+}\left(\frac{m+M}{2}\right)\right]$ provided $x_i \in [m,M]$ and $w_i \geq 0$ $(i = 1,\ldots,n)$ with $W_n := \sum_{i=1}^{n} w_i = 1$.

Moreover, if we assume that $\Phi : [m,M] \to \mathbb{R}$ is differentiable on (m,M), then

$$\begin{aligned}
&\sum_{i=1}^{n} w_i \Phi(x_i) \qquad (2.23)\\
&\leq \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \sum_{i=1}^{n} w_i \Phi'(x_i) \sum_{i=1}^{n} w_i \left(x_i - \frac{m+M}{2} \right)\\
&\quad + \sum_{i=1}^{n} w_i x_i \left[\Phi'(x_i) - \sum_{j=1}^{n} w_j \Phi'(x_j) \right]\\
&\leq \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \sum_{i=1}^{n} w_i \Phi'(x_i) \sum_{i=1}^{n} w_i \left(x_i - \frac{m+M}{2} \right)\\
&\quad + \frac{1}{2}\left[\Phi'_{+}(M) - \Phi'_{-}(m)\right] \sum_{i=1}^{n} w_i \left| x_i - \sum_{j=1}^{n} w_j \Phi'(x_j) \right|\\
&\leq \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \sum_{i=1}^{n} w_i \Phi'(x_i) \sum_{i=1}^{n} w_i \left(x_i - \frac{m+M}{2} \right)\\
&\quad + \frac{1}{2}\left[\Phi'_{+}(M) - \Phi'_{-}(m)\right] \left[\sum_{i=1}^{n} w_i x_i^2 - \left(\sum_{i=1}^{n} w_i x_i \right)^2 \right]^{\frac{1}{2}}\\
&\leq \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \sum_{i=1}^{n} w_i \Phi'(x_i) \sum_{i=1}^{n} w_i \left(x_i - \frac{m+M}{2} \right)\\
&\quad + \frac{1}{4}\left[\Phi'_{+}(M) - \Phi'_{-}(m)\right](M-m),
\end{aligned}$$

provided $x_i \in [m,M]$ and $w_i \geq 0$ $(i = 1,\ldots,n)$ with $W_n := \sum_{i=1}^{n} w_i = 1$.

If we assume that

$$\sum_{i=1}^{n} w_i x_i = \frac{m+M}{2}$$

then from (2.22) and (2.23) we have

$$\Phi\left(\frac{m+M}{2}\right) \leq \sum_{i=1}^{n} w_i \Phi(x_i) \leq \frac{\Phi(m)+\Phi(M)}{2} \tag{2.24}$$

and

$$\begin{aligned}
&\sum_{i=1}^{n} w_i \Phi(x_i) \qquad (2.25)\\
&\leq \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \sum_{i=1}^{n} w_i x_i \left[\Phi'(x_i) - \sum_{j=1}^{n} w_j \Phi'(x_j)\right]\\
&\leq \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \frac{1}{2}\left[\Phi'_+(M) - \Phi'_-(m)\right]\sum_{i=1}^{n} w_i \left| x_i - \sum_{j=1}^{n} w_j \Phi'(x_j)\right|\\
&\leq \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \frac{1}{2}\left[\Phi'_+(M) - \Phi'_-(m)\right] \times \left[\sum_{i=1}^{n} w_i x_i^2 - \left(\sum_{i=1}^{n} w_i x_i\right)^2\right]^{\frac{1}{2}}\\
&\leq \frac{1}{M-m}\int_m^M \Phi(t)\,dt + \frac{1}{4}\left[\Phi'_+(M) - \Phi'_-(m)\right](M-m).
\end{aligned}$$

3. APPLICATIONS FOR f-DIVERGENCE

One of the important issues in many applications of Probability Theory is finding an appropriate measure of *distance* (or *difference* or *discrimination*) between two probability distributions. A number of divergence measures for this purpose have been proposed and extensively studied by Jeffreys [24], Kullback and Leibler [30], Rényi [40], Havrda and Charvat [21], Kapur [27], Sharma and Mittal [42], Burbea and Rao [7], Rao [39], Lin [31], Csiszár [11], Ali and Silvey [3], Vajda [50], Shioya and Da-te [16] and others (see for example [32] and the references therein).

These measures have been applied in a variety of fields such as: anthropology [39], genetics [32], finance, economics, and political science [41,47,48], biology [37], the analysis of contingency tables [20], approximation of probability distributions [9,28], signal processing [25,26] and pattern recognition [5,8]. A number of these measures of distance are specific cases of Csiszár f-divergence and so further exploration of this concept will have a flow on effect to other measures of distance and to areas in which they are applied.

Assume that a set Ω and the σ-finite measure μ are given. Consider the set of all probability densities on μ to be $\mathcal{P} := \left\{p \,|\, p : \Omega \to \mathbb{R},\ p(x) \geq 0,\ \int_\Omega p(x)\,d\mu(x) = 1\right\}$.

The Kullback-Leibler divergence [30] is well known among the information divergences. It is defined as:

$$D_{KL}(p,q) := \int_{\Omega} p(x) \ln\left[\frac{p(x)}{q(x)}\right] d\mu(x), \quad p,q \in \mathcal{P}, \tag{3.1}$$

where ln is to base e.

In Information Theory and Statistics, various divergences are applied in addition to the Kullback-Leibler divergence. These are the: *variation distance* D_v, *Hellinger distance* D_H [22], χ^2*-divergence* D_{χ^2}, α*-divergence* D_α, *Bhattacharyya distance* D_B [6], *Harmonic distance* D_{Ha}, *Jeffrey's distance* D_J [24], *triangular discrimination* D_Δ [49], etc... They are defined as follows:

$$D_v(p,q) := \int_{\Omega} |p(x) - q(x)| \, d\mu(x), \quad p,q \in \mathcal{P}; \tag{3.2}$$

$$D_H(p,q) := \int_{\Omega} \left|\sqrt{p(x)} - \sqrt{q(x)}\right| d\mu(x), \quad p,q \in \mathcal{P}; \tag{3.3}$$

$$D_{\chi^2}(p,q) := \int_{\Omega} p(x)\left[\left(\frac{q(x)}{p(x)}\right)^2 - 1\right] d\mu(x), \quad p,q \in \mathcal{P}; \tag{3.4}$$

$$D_\alpha(p,q) := \frac{4}{1-\alpha^2}\left[1 - \int_{\Omega} [p(x)]^{\frac{1-\alpha}{2}} [q(x)]^{\frac{1+\alpha}{2}} d\mu(x)\right], \quad p,q \in \mathcal{P}; \tag{3.5}$$

$$D_B(p,q) := \int_{\Omega} \sqrt{p(x)q(x)} \, d\mu(x), \quad p,q \in \mathcal{P}; \tag{3.6}$$

$$D_{Ha}(p,q) := \int_{\Omega} \frac{2p(x)q(x)}{p(x)+q(x)} d\mu(x), \quad p,q \in \mathcal{P}; \tag{3.7}$$

$$D_J(p,q) := \int_{\Omega} [p(x) - q(x)] \ln\left[\frac{p(x)}{q(x)}\right] d\mu(x), \quad p,q \in \mathcal{P}; \tag{3.8}$$

$$D_\Delta(p,q) := \int_{\Omega} \frac{[p(x) - q(x)]^2}{p(x)+q(x)} d\mu(x), \quad p,q \in \mathcal{P}. \tag{3.9}$$

For other divergence measures, see the paper [27] by Kapur or the book on line [46] by Taneja.

Csiszár f-divergence is defined as follows [12]

$$I_f(p,q) := \int_{\Omega} p(x) f\left[\frac{q(x)}{p(x)}\right] d\mu(x), \quad p,q \in \mathcal{P}, \tag{3.10}$$

where f is convex on $(0,\infty)$. It is assumed that $f(1) = 0$ and strictly convex around 1. By appropriately defining this convex function, various divergences are derived. Most of the above distances (3.1) – (3.9), are particular instances of Csiszár f-divergence. There are also many others which are not in this class (see for example [46]).

For the basic properties of Csiszár f-divergence see [12, 13] and [50]. The most important property is that

$$I_f(p,q) \geq 0 \text{ for any } p,q \in \mathcal{P}.$$

The definition (3.10) can be extended for any measurable function defined on $[0,\infty)$.

Proposition 1. *Let $f:(0,\infty)\to\mathbb{R}$ be a convex function with the property that $f(1)=0$. Assume that $p,q\in\mathcal{P}$ and there exists the constants $0<r<1<R<\infty$ such that*

$$r \leq \frac{q(x)}{p(x)} \leq R \text{ for } \mu\text{-a.e. } x \in \Omega. \tag{3.11}$$

Then we have the inequalities

$$\begin{aligned} & f\left(\frac{r+R}{2}\right)+\varphi\left(\frac{r+R}{2}\right)\left(1-\frac{r+R}{2}\right) \\ & \leq I_f(p,q) \\ & \leq \frac{f(r)+f(R)}{2}+\frac{f(R)-f(r)}{R-r}\left(1-\frac{r+R}{2}\right), \end{aligned} \tag{3.12}$$

where $\varphi\left(\frac{r+R}{2}\right)\in\left[f'_-\left(\frac{r+R}{2}\right),f'_+\left(\frac{r+R}{2}\right)\right]$.

Proof. From the inequality (2.1) applied for the convex function f we have

$$\begin{aligned} & f\left(\frac{r+R}{2}\right)+\varphi\left(\frac{r+R}{2}\right)\int_\Omega\left(\frac{q(x)}{p(x)}-\frac{r+R}{2}\right)p(x)\,d\mu(x) \\ & \leq \int_\Omega f\left(\frac{q(x)}{p(x)}\right)p(x)\,d\mu(x) \\ & \leq \frac{f(r)+f(R)}{2}+\frac{f(R)-f(r)}{R-r}\int_\Omega\left(\frac{q(x)}{p(x)}-\frac{r+R}{2}\right)p(x)\,d\mu(x), \end{aligned} \tag{3.13}$$

where $\varphi\left(\frac{r+R}{2}\right)\in\left[f'_-\left(\frac{r+R}{2}\right),f'_+\left(\frac{r+R}{2}\right)\right]$.

Since

$$\int_\Omega\left(\frac{q(x)}{p(x)}-\frac{r+R}{2}\right)p(x)\,d\mu(x)=1-\frac{r+R}{2},$$

then by (3.13) we get the desired result (3.12). □

We also have:

Proposition 2. *Let $f:(0,\infty)\to\mathbb{R}$ be a differentiable convex function with the property that $f(1)=0$. Assume that $p,q\in\mathcal{P}$ and there exists the constants $0<r<1<R<\infty$ such that (3.11) is valid. Then*

$$\frac{2}{R-r}\int_r^R f(s)\,ds-\frac{f(R)+f(r)}{2}+\frac{f(R)-f(r)}{R-r}\left(1-\frac{r+R}{2}\right) \tag{3.14}$$

$$
\begin{aligned}
&\le I_f(p,q) \\
&\le \frac{1}{R-r}\int_r^R f(s)\,ds + I_{f'}(p,q)\left(1-\frac{r+R}{2}\right) + I_{f'\ell}(p,q) - I_{f'}(p,q) \\
&\le \frac{1}{R-r}\int_r^R f(s)\,ds + I_{f'}(p,q)\left(1-\frac{r+R}{2}\right) \\
&\quad + \frac{1}{2}\left[f'_+(R) - f'_-(r)\right] D_v(p,q) \\
&\le \frac{1}{R-r}\int_r^R f(s)\,ds + I_{f'}(p,q)\left(1-\frac{r+R}{2}\right) \\
&\quad + \frac{1}{2}\left[f'_+(R) - f'_-(r)\right] D_{\chi^2}^{1/2}(p,q) \\
&\le \frac{1}{R-r}\int_r^R f(s)\,ds + I_{f'}(p,q)\left(1-\frac{r+R}{2}\right) \\
&\quad + \frac{1}{4}\left[f'_+(R) - f'_-(r)\right](R-r),
\end{aligned}
$$

where $\ell(t) = t$, $t \in \mathbb{R}$.

Proof. From the inequalities (2.7) and (2.14) for the convex function f, we have

$$
\begin{aligned}
&\frac{2}{R-r}\int_r^R f(s)\,ds - \frac{f(R)+f(r)}{2} + \frac{f(R)-f(r)}{R-r}\int_\Omega \left(\frac{q(x)}{p(x)} - \frac{r+R}{2}\right) p(x)\,d\mu(x) \\
&\le \int_\Omega f\left(\frac{q(x)}{p(x)}\right) p(x)\,d\mu(x) \\
&\le \frac{1}{R-r}\int_r^R f(s)\,ds + \int_\Omega f'\left(\frac{q(x)}{p(x)}\right) p(x)\,d\mu(x) \times \int_\Omega \left(\frac{q(x)}{p(x)} - \frac{r+R}{2}\right) p(x)\,d\mu(x) \\
&\quad + \int_\Omega f'\left(\frac{q(x)}{p(x)}\right) \frac{q(x)}{p(x)} p(x)\,d\mu(x) - \int_\Omega f'\left(\frac{q(x)}{p(x)}\right) p(x)\,d\mu(x) \int_\Omega \frac{q(x)}{p(x)} p(x)\,d\mu(x) \\
&\le \frac{1}{R-r}\int_r^R f(s)\,ds + \int_\Omega f'\left(\frac{q(x)}{p(x)}\right) p(x)\,d\mu(x) \int_\Omega \left(\frac{q(x)}{p(x)} - \frac{r+R}{2}\right) p(x)\,d\mu(x) \\
&\quad + \frac{1}{2}\left[f'_+(R) - f'_-(r)\right] \int_\Omega \left|\frac{q(x)}{p(x)} - \int_\Omega \frac{q(y)}{p(y)} p(y)\,d\mu(y)\right| p(x)\,d\mu(x) \\
&\le \frac{1}{R-r}\int_r^R f(s)\,ds + \int_\Omega f'\left(\frac{q(x)}{p(x)}\right) p(x)\,d\mu(x) \int_\Omega \left(\frac{q(x)}{p(x)} - \frac{r+R}{2}\right) p(x)\,d\mu(x) \\
&\quad + \frac{1}{2}\left[f'_+(R) - f'_-(r)\right] \left[\int_\Omega \left(\frac{q(x)}{p(x)}\right)^2 p(x)\,d\mu(x) - \left(\int_\Omega \frac{q(x)}{p(x)} p(x)\,d\mu(x)\right)^2\right]^{\frac{1}{2}} \\
&\le \frac{1}{R-r}\int_r^R f(s)\,ds + \int_\Omega f'\left(\frac{q(x)}{p(x)}\right) p(x)\,d\mu(x) \int_\Omega \left(\frac{q(x)}{p(x)} - \frac{r+R}{2}\right) p(x)\,d\mu(x)
\end{aligned}
$$

$$+\frac{1}{4}\left[f'_+(R)-f'_-(r)\right](R-r), \tag{3.15}$$

which is equivalent to the desired result (3.14). □

Consider the convex function $f(t)=-\ln t$, $t>0$. Then

$$\begin{aligned} I_{-\ln}(p,q) &:= -\int_\Omega p(x)\ln\left[\frac{q(x)}{p(x)}\right]d\mu(x)=\int_\Omega p(x)\ln\left[\frac{p(x)}{q(x)}\right]d\mu(x) \\ &= D_{KL}(p,q), \quad p,q\in\mathcal{P} \end{aligned}$$

and

$$\begin{aligned} &\frac{f(r)+f(R)}{2}+\frac{f(R)-f(r)}{R-r}\left(1-\frac{r+R}{2}\right) \\ &= -\frac{\ln r+\ln R}{2}+\frac{-\ln R+\ln r}{R-r}\left(1-\frac{r+R}{2}\right) \\ &= \frac{\ln R-\ln r}{R-r}\left(\frac{r+R}{2}-1\right)-\frac{\ln r+\ln R}{2} \\ &= \ln\left(\frac{R}{r}\right)^{\frac{\frac{r+R}{2}-1}{R-r}}-\ln\sqrt{rR}=\ln\left(\frac{\left(\frac{R}{r}\right)^{\frac{\frac{r+R}{2}-1}{R-r}}}{\sqrt{rR}}\right) \end{aligned}$$

and by the second inequality in (3.12) we get

$$(0\le)\, D_{KL}(p,q)\le\ln\left(\frac{\left(\frac{R}{r}\right)^{\frac{\frac{r+R}{2}-1}{R-r}}}{\sqrt{rR}}\right). \tag{3.16}$$

We also have that

$$\begin{aligned} &\frac{1}{R-r}\int_r^R f(s)\,ds+I_{f'}(p,q)\left(1-\frac{r+R}{2}\right)+\frac{1}{2}\left[f'_+(R)-f'_-(r)\right]D_v(p,q) \\ &= -\frac{1}{R-r}\int_r^R \ln s\,ds-\int_\Omega p(x)\left[\frac{p(x)}{q(x)}\right]d\mu(x)\left(1-\frac{r+R}{2}\right) \\ &\quad+\frac{1}{2}\left(-\frac{1}{R}+\frac{1}{r}\right)D_v(p,q) \\ &= -\ln I(r,R)-\left(D_{\chi^2}(p,q)+1\right)\left(1-\frac{r+R}{2}\right)+\frac{1}{2}\frac{R-r}{rR}D_v(p,q) \\ &= \frac{1}{2}\frac{R-r}{rR}D_v(p,q)+\left(D_{\chi^2}(p,q)+1\right)\left(\frac{r+R}{2}-1\right)-\ln I(r,R) \end{aligned}$$

and by the third inequality in (3.14) we have

$$
\begin{aligned}
&(0 \leq) D_{KL}(p,q) \\
&\leq \frac{R-r}{2rR} D_v(p,q) + \left(D_{\chi^2}(p,q)+1\right)\left(\frac{r+R}{2}-1\right) - \ln I(r,R),
\end{aligned} \tag{3.17}
$$

where $I(a,b)$ is the *identric mean*, namely

$$
\begin{aligned}
I(a,b) &:= \begin{cases} \frac{1}{e}\cdot\left(\frac{b^b}{a^a}\right)^{\frac{1}{b-a}} & \text{if } a \neq b \\ a & \text{if } a = b \end{cases} \\
&= \exp\left(\frac{1}{b-a}\int_a^b \ln s\, ds\right), \text{ if } a \neq b.
\end{aligned}
$$

References

[1] S. Abtamovich, "Convexity, subadditivity and generalized Jensen's inequality," *Ann. Funct. Anal.*, vol. 4, no. 2, pp. 183–194, 2013.

[2] J. M. Adaz, "A measure-theoretic version of the dragomir-jensen inequality," *Proc. Amer. Math. Soc.*, vol. 140, no. 7, pp. 2391–2399, 2012.

[3] S. M. Ali and S. D. Silvey, "A general class of coefficients of divergence of one distribution from another," *J. Roy. Statist. Soc. Sec B*, vol. 28, no. 1, pp. 131–142, 1966.

[4] F. Bahrami, E. B. A., and S. M. Manjegani, "Operator-valued bochner integrable functions and jensen's inequality," *Georgian Math. J.*, vol. 20, no. 4, pp. 625–640, 2013.

[5] M. Ben-Bassat, "f-entropies, probability of error and feature selection," *Inform. Control*, vol. 39, pp. 227–242, 1978, doi: 10.1016/S0019-9958(78)90587-9.

[6] A. Bhattacharyya, "On a measure of divergence between two statistical populations defined by their probability distributions," *Bull. Calcutta Math. Soc.*, vol. 35, no. 1, pp. 99–109, 1943.

[7] I. Burbea and C. R. Rao, "On the convexity of some divergence measures based on entropy function," *IEEE Trans. Inf. Th.*, vol. 28, no. 3, pp. 489–495, 1982.

[8] C. H. Chen, *Statistical Pattern Recognition*. Rocelle Park, New York: Hoyderc Book Co.r, 1973.

[9] C. K. Chow and C. N. Lin, "Approximating discrete probability distributions with dependence trees," *IEEE Trans. Inf. Th.*, vol. 14, no. 3, pp. 462–467, 1968, doi: 10.1109/TIT.1968.1054142.

[10] V. Cirtoaje, "An extension of jensen's discrete inequality to partially convex functions," *J. Inequal. Appl.*, vol. 2013, no. 2013:54, pp. 1–23, 2013, doi: 10.1186/1029-242X-2013-54.

[11] I. Csiszár, "Information-type measures of difference of probability distributions and indirect observations," *Studia Math. Hungarica*, vol. 2, pp. 299–318, 1967.

[12] I. Csiszár, "On topological properties of $f-$divergences," *Studia Math. Hungarica*, vol. 2, pp. 329–339, 1967.

[13] I. Csiszár and J. Körner, *Information Theory: Coding Theorem for Discrete Memoryless Systems*. New York,: Academic Press, 1981.

[14] S. S. Dragomir, "On a reverse of jessen's inequality for isotonic linear functionals," *J. Ineq. Pure & Appl. Math.*, vol. 2, no. 3, p. Art. 36, 2001.

[15] S. S. Dragomir, "A grüss type inequality for isotonic linear functionals and applications," *Demonstratio Math*, vol. 36, no. 3, pp. 551–562, 2003.

[16] S. S. Dragomir, "Bounds for the deviation of a function from the chord generated by its extremities," *Bull. Aust. Math. Soc.*, vol. 78, no. 2, pp. 225–248, 2008, doi: 10.1017/S0004972708000671.

[17] S. S. Dragomir and N. M. Ionescu, "Some converse of jensen's inequality and applications," *Rev. Anal. Numér. Théor. Approx.*, vol. 23, no. 1, pp. 71–78, 1994.

[18] A. El Frisi, B. Belaidi, and Z. Lareuch, "Jensen type inequalities for twice differentiable functions," *J. Nonlinear Sci. Appl.*, vol. 5, no. 5, pp. 350–356, 2012.

[19] L. Fejér, "Über die fourierreihen, ii," *Anz Ungar. Akad. Wiss.*, vol. 24, pp. 369–390, 1906.

[20] D. Gokhale and S. Kullback, *Information in Contingency Tables*. New York: Marcel Decker, 1978.

[21] J. H. Havrda and F. Charvat, "Quantification method classification process: concept of structural α-entropy," *Kybernetika*, vol. 3, pp. 210–271, 1967.

[22] E. Hellinger, "Neue bergrüirdung du theorie quadratisher formerus von uneudlichvieleu veränderlicher," *J. für reine and Augeur. Math.*, vol. 36, pp. 210–271, 1909.

[23] L. Horváth, "A refinement of the integral form of jensen's inequality," *J. Inequal. Appl.*, vol. 3, no. 2012:178, pp. 1–19, 2012, doi: 10.1186/1029-242X-2012-178.

[24] H. Jeffreys, "An invariant form for the prior probability in estimating problems," *Proc. Roy. Soc. London*, vol. 186, pp. 453–461, 1946, doi: 10.1098/rspa.1946.0056.

[25] T. T. Kadota and L. A. Shepp, "On the best finite set of linear observables for discriminating two gaussian signals," *IEEE Trans. Inf. Th.*, vol. 13, pp. 288–294, 1967, doi: 10.1109/TIT.1967.1054013.

[26] T. Kailath, "The divergence and bhattacharyya distance measures in signal selection," *IEEE Trans. Comm. Technology*, vol. 15, pp. 52–60, 1967, doi: 10.1109/TCOM.1967.1089532.

[27] J. N. Kapur, "A comparative assessment of various measures of directed divergence," *Advances in Management Studies*, vol. 3, pp. 1–16, 1984.

[28] D. Kazakos and T. Cotsidas, "A decision theory approach to the approximation of discrete probability densities,," *IEEE Trans. Perform. Anal. Machine Intell.*, vol. 1, pp. 61–67, 1980, doi: 10.1109/TPAMI.1980.4766971.

[29] M. A. Khan, G. A. Khan, T. Ali, T. Batbold, and A. Kiliçman, "Further refinements of jensen's type inequalities for the function defined on the rectangle," *Abstr. Appl. Anal.*, vol. 2013, no. Art. ID 214123, pp. 1–8, 2013.

[30] S. Kullback and R. A. Leibler, "On information and sufficiency," *Annals Math. Statist.*, vol. 22, pp. 79–86, 1951.

[31] J. Lin, "Divergence measures based on the shannon entropy," *IEEE Trans. Inf. Th.*, vol. 37, no. 1, pp. 145–151, 1991, doi: 10.1109/18.61115.

[32] M. Mei, "The theory of genetic distance and evaluation of human races," *Japan J. Human Genetics*, vol. 23, pp. 341–369, 1978.

[33] M. V. Mihai, "Jensen's inequality for mixed convex functions of two real variables," *Acta Univ. Apulensis Math. Inform.*, vol. 34, pp. 179–183, 2013.

[34] F. C. Mitroi, "On the jensen-steffensen inequality and superquadraticity," *An. Univ. Oradea Fasc. Mat.*, vol. 18, no. 1, pp. 269–275, 2011.

[35] M. S. Moslehian and M. Kian, "Jensen type inequalities for q-class functions," *Bull. Aust. Math. Soc.*, vol. 85, no. 1, pp. 128–142, 2012, doi: 10.1017/S0004972711002863.

[36] Z. Pavić, "The applications of functional variants of jensen's inequality," *J. Funct. Spaces Appl.*, vol. 2013, no. Art. ID 194830, pp. 1–5, 2013, doi: 10.1155/2013/194830.

[37] E. C. Pielou, *Ecological Diversity*. New York: Wiley, 1975.

[38] F. Popovici and C.-I. Spitidon, "The jensen inequality for (m,n)-convex functions," *An. Univ. Craiova Ser. Mat. Inform.*, vol. 38, no. 4, pp. 63–66, 2011.

[39] C. R. Rao, "Diversity and dissimilarity coefficients: a unified approach," *Theoretic Population Biology*, vol. 21, pp. 24–43, 1982, doi: 10.1016/0040-5809(82)90004-1.

[40] A. Rényi, "On measures of entropy and information," *Proc. Fourth Berkeley Symp. Math. Stat. and Prob., University of California Press*, vol. 1, pp. 547–561, 1961.

[41] A. Sen, *On Economic Inequality*. London: Oxford University Press, 1973.
[42] B. D. Sharma and D. P. Mittal, "New non-additive measures of relative information," *Journ. Comb. Inf. Sys. Sci.*, vol. 2, no. 4, pp. 22–132, 1977.
[43] R. Sharma, "On jensen's inequality for positive linear functionals," *Int. J. Math. Sci. Eng. Appl.*, vol. 5, no. 5, pp. 263–271, 2011.
[44] S. Simic, "Sharp global bounds for jensen's inequality," *Rocky Mountain J. Math.*, vol. 41, no. 6, pp. 2021–2031, 2011, doi: 10.1216/RMJ-2011-41-6-2021.
[45] H. M. Srivastava, Z.-H. Zhang, and Y.-D. Wu, "Some further refinements and extensions of the hermite-hadamard and jensen inequalities in several variables," *Math. Comput. Modelling*, vol. 54, no. 11-12, pp. 2709–2717, 2011, doi: 10.1016/j.mcm.2011.06.057.
[46] I. J. Taneja, *Generalised information measures and their applications*. http://www.mtm.ufsc.br/˜taneja/bhtml/bhtml.html: Book on line, 2001.
[47] H. Theil, *Economics and Information Theory*. Amsterdam: North-Holland, 1972.
[48] H. Theil, *Statistical Decomposition Analysis*. Amsterdam: North-Holland, 1972.
[49] F. Topsoe, "Some inequalities for information divergence and related measures of discrimination," *Res. Rep. Coll., RGMIA,*, vol. 2, no. 1, pp. 85–98, 1999.
[50] I. Vajda, *Theory of Statistical Inference and Information*. Dordrecht-Boston: Kluwer Academic Publishers, 1989.
[51] G. Zabandan and A. Kiliçman, "A new version of Jensen's inequality and related results," *J. Inequal. Appl.*, vol. 2012, no. 2012:238, pp. 1–7, 2012, doi: 10.1186/1029-242X-2012-238.

Author's address

S. S. Dragomir
Mathematics, School of Engineering & Science, Victoria University, PO Box 14428, Melbourne City, MC 8001, Australia,
School of Computational & Applied Mathematics,, University of the Witwatersrand,, Private Bag 3, Johannesburg 2050, South Africa
E-mail address: sever.dragomir@vu.edu.au